

OPTICAL CHARACTER READER**RECEIVED**

DEC 20 2004

Technology Center 2600

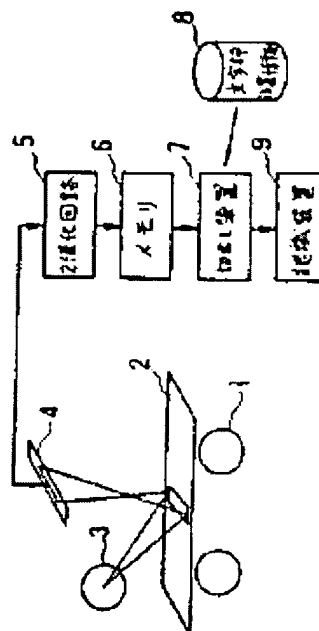
Patent number: JP1028792
Publication date: 1989-01-31
Inventor: UCHIYAMA HIROMI; others: 04
Applicant: HITACHI LTD; others: 01
Classification:
- international: G06K9/34
- european:
Application number: JP19870183417 19870724
Priority number(s):

Abstract of JP1028792

PURPOSE:To easily detect a character entering frame, to remove it and to easily extract an entering character pattern by providing a means for searching the contour point of an image pattern including the character entering frame, a deciding means and a removing means.

CONSTITUTION:A document 2 is carried by a document carrying mechanism 1, scanned by a light source 3 at the time of reaching a reading part consisting of a light emitting source 3 and a photodetector 4, the reflected light of the document 2 is incident on the element 4 and the detector 4 converts the frame pattern and the character pattern on the document 2 to an electric signal and outputs. A binarization circuit 5 converts the input electric signal from the element 4 to the binary signal of 0 or 1 according to a prescribed threshold and outputs to a memory 6. The memory 6 stores the image or the pattern on the document 2 as a binary signal pattern.

Thereby, a segmenting device 7 refers to character frame position information stored in a file device 8, traces the contour to easily detect the character entering frame from the pattern in the memory 6, remove it, extract only the character pattern and output respective one character to a recognition device 9.



Data supplied from the **esp@cenet** database - Patent Abstracts of Japan

BEST AVAILABLE COPY

⑫ 公開特許公報(A)

昭64-28792

⑮ Int.Cl.⁴

識別記号

庁内整理番号

⑬ 公開 昭和64年(1989)1月31日

G 06 K 9/34

6942-5B

審査請求 未請求 発明の数 1 (全4頁)

⑭ 発明の名称 光学文字読取装置

⑰ 特 願 昭62-183417

⑱ 出 願 昭62(1987)7月24日

⑲ 発 明 者 内 山 広 美 神奈川県小田原市国府津2880番地 日立コンピュータ機器株式会社内

⑳ 発 明 者 吉 田 昭 信 神奈川県小田原市国府津2880番地 日立コンピュータ機器株式会社内

㉑ 発 明 者 木 下 市 朗 神奈川県小田原市国府津2880番地 日立コンピュータ機器株式会社内

㉒ 出 願 人 株式会社日立製作所 東京都千代田区神田駿河台4丁目6番地

㉓ 出 願 人 日立コンピュータ機器株式会社 神奈川県小田原市国府津2880番地

㉔ 代 理 人 弁理士 中村 純之助
最終頁に続く

明 細 書

帳票を読取る光学文字読取装置に関する。

〔従来の技術〕

1. 発明の名称

光学文字読取装置

2. 特許請求の範囲

1. 文字パターンより小さい所定の大きさのドットよりなる文字記入枠を備えた帳票を読取る光学文字読取装置において、文字記入枠を含むイメージパターンの輪郭点を探索する手段と、前記探索により発見された輪郭点を持つパターンの輪郭のサイズ情報により文字記入枠の輪郭であるか否かを判定する手段と、文字記入枠のパターンを除去する手段を具備し、記入文字パターンのみを抽出することを特徴とする光学文字読取装置。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は、文字記入枠を備えた帳票の文字を読取る光学文字読取装置に係り、特に、所定の大きさのドットよりなる文字記入枠パターンを備えた

所定の大きさのドットで印刷された文字記入枠を備えた帳票の文字を読取る光学式文字読取装置は、特公昭61-30313号に記載のように次のような文字分離を行っていた。すなわち量子化パターンの最も左の縦走査列位置から始めて右の方へ移動しながら点検出(一定範囲以下の大きさを持つ信号“1”のかたまりを点とする。)により検出された点が一定個数以上に達した走査列位置を左端とする。左端に文字記入枠の幅を加えた走査列位置から縦の点線の位置ずれ幅を加え、または減じることによって得られる縦の点線の存在すべき範囲において、検出された点が一定個数以上に達した走査列位置を右端とする。これにより左端と右端の間が1つの枠と見なせる。更に、枠の右端から隣接した枠の間隔および縦の点線の位置ずれ幅を加えることによって得られる縦の点線の存在すべき範囲において、前記点検出により次の枠の左端を検出する。更に、同様の走査を行って

右端を検出し、検出された点をそれぞれ消去する。更に、文字記入枠の上端及び下端の最初の横の点に対しては前記縦の点線の存在すべき範囲内において文字枠の上端及び下端の存在すべき範囲を設定し、前記範囲内において点検出し、検出された点の最も上の点の位置と最も下の点の位置をそれぞれ上位の点位置、下位の点位置とし前記上位の点位置には横の点間の位置ずれ幅を加え、前記下位の点位置には横の点間の位置ずれ幅を減じることにより文字枠の上端及び下端の存在すべき範囲を決定し、この範囲内において検出された点を消去し、文字パターンをのみ検出を行っていた。
〔発明が解決しようとする問題点〕

上記従来技術は、枠の左端を検出し、その左端位置より枠位置ずれを考慮し右端枠検出の範囲設定を行ない、右端の枠の検出を行なう、上端及び下端の枠は、左右端枠の最上位点、最下位点から検出範囲を設定し、検出する。次の枠検出は、右端位置より枠間の情報と枠位置ずれを考慮して検出範囲の設定により、次の枠を検出する。このよ

うに枠の仮想位置情報により範囲設定を行なっているため、行間、枠間が小さい場合や枠のそばに文字等が記入された場合に、範囲設定がむずかしく誤読又は検出不能となる等の問題があった。

本発明の目的は、容易に文字記入枠を検出し、これを除去して記入文字パターンを容易に抽出し得る光学文字読取装置を提供することにある。

〔問題点を解決するための手段〕

上記目的は、文字パターンより小さい所定の大きさのドットよりなる文字記入枠を備えた帳票を読取る光学文字読取装置において、文字記入枠を含むイメージパターンの輪郭点を探索する手段と、前記探索により発見された輪郭点を持つパターンの輪郭のサイズ情報により文字記入枠の輪郭であるか否かを判定する手段と、文字記入枠のパターンを除去する手段を具備し、記入文字パターンのみを抽出することを特徴とする光学文字読取装置により達成される。

イメージパターンの輪郭点を探索する手段及び探索された輪郭点を持つパターンの輪郭を追跡す

る方法は公知の方法を用いればよく、例えば特開昭61-36878号の記載の方法で行なえばよい。

〔作 用〕

文字記入枠の輪郭か否かの判定手段は、発見された輪郭のサイズ情報を輪郭閾値と比較判定する。そのため枠に隣接して文字が記載してあったり、また枠間が小さい場合でも誤読することがない。
〔実施例〕

以下、添付の図面に示す実施例により、更に詳細に本発明について説明する。

第1図は、本発明が適用された光学文字読取装置の構成の一例を示す。

第1図において、1は搬送ローラを含む帳票搬送機構、2は帳票、3は発光源、4は受光素子、5は2値化回路、6はメモリ、7は切出し装置、8はファイル装置で文字枠位置情報が格納されている。また9は認識装置である。

帳票2は、帳票搬送機構1で搬送され、発光源3及び受光素子4からなる読取部に達すると、発光源3により光学的に走査される。帳票2で反射

された光が受光素子4に入射し、受光素子4は、順次入射される帳票上の枠パターンおよび文字パターンを電気信号に変換し出力する。受光素子4は、半導体CCDセンサ等が用いられる。2値化回路5は、入力電気信号を所定の閾値により、“0”又は“1”の2値信号に変換してメモリ6に出力する。メモリ6は、帳票上のパターンの像を2値信号パターンとして格納する。切出し装置7は、ファイル装置8に格納されている文字枠位置情報を参照するとともに、輪郭追跡を行なってメモリ6内のパターンから文字パターンのみを抽出し、抽出した文字パターンを1文字ずつ認識装置9に出力する。

次に、切出し装置7における文字切出し処理の詳細を説明する。

帳票から読取られてメモリ6に格納されたままのパターンは、第2図と同じように枠パターン $f_1 \sim f_3$ と文字パターン $c_1 \sim c_3$ からなる。図中Wは枠幅、Hは枠の高さ、gは枠の間隔である。切出し装置7は、ファイル装置8に格納されてい

る枠パターン f_1 の位置情報 x 、 y 、 W 、 H 、 g を読み出して、枠パターン f_1 より g_1 、 g_2 、 g_3 だけ外側にウィンドウ a を設定する。本実施例の場合、 g_1 は約2mm、 g_2 は約3mm、 g_3 は $1.5W + g - g_1$ である。前記設定されたウィンドウ a 内で、パターン f_1 の輪郭点をサーチし、ウィンドウ a 内に存在するパターンについて輪郭を追跡して、そのパターンの全輪郭を第4図C1'とD1~Dnのように抽出する。前記の輪郭パターン抽出の方法は、特開昭61-36878に記載する方法により容易にできる。抽出された輪郭サイズ情報の輪郭長さ、高さ、幅を文字記入枠輪郭の各閾値と比較し、小さな輪郭と判定されたパターンを除去する。前記を繰返すことにより、文字記入枠D1~Dnを除去し、除去できなかったパターンを記入文字パターンとし、第5図に示す成分リストテーブルを作成することにより、記入文字だけを容易に抽出できる。

切出し装置7は、第4図の文字パターンC1'の輪郭パターンデータを認識装置9に転送するこ

とにより、文字の認識ができる。

(発明の効果)

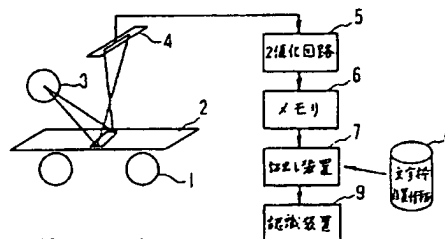
以上の説明から明らかな様に、本発明によれば文字記入枠パターンと記入文字パターンを容易に分離、抽出することができるので、文字と枠が同色でも読取りが可能になり、帳票設計に対する制約が緩和された。

4. 図面の簡単な説明

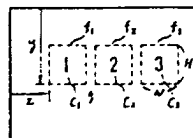
第1図は、本発明の一実施例を示すブロック図、第2図は、本発明の処理の対象となるイメージパターンの一例を示す図、第3図は、第2図のパターンから本発明により設定されるウィンドウの一例を示す図、第4図は、ウィンドウ内の輪郭パターンを示す図、第5図は、抽出されたパターンの記録するテーブル構成図である。

- | | |
|----------|----------|
| 1…帳票搬送機構 | 2…帳票 |
| 3…発光源 | 4…受光素子 |
| 5…2値化回路 | 6…メモリ |
| 7…切出し装置 | 8…ファイル装置 |
| 9…認識装置 | |

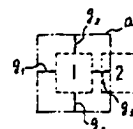
第1図



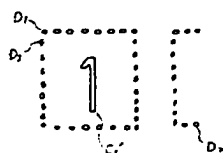
第2図



第3図



第4図



第5図

1	2	3	4	5	6	7	8	9

- 1…帳票搬送機構
2…帳票
3…発光源
4…受光素子

第1頁の続き

⑦発 明 者 上 田 欣 洋 神奈川県小田原市国府津2880番地 日立コンピュータ機器
株式会社内
⑧発 明 者 栗 野 清 道 神奈川県小田原市国府津2880番地 株式会社日立製作所小
田原工場内

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.